

Un autre regard sur l'apprenant avancé : sa capacité de paraphraser

Alexandra Tsedryk

Mount Saint Vincent University, Canada

Alexandra.tsedryk@msvu.ca

1 Compétence paraphrastique de l'apprenant avancé

Savoir paraphraser, c'est savoir reformuler des phrases en créant des énoncés sémantiquement équivalents, comme ceux en (1).

- (1)
- a. *Julie admire énormément l'auteur de ce roman historique.*
 - b. *L'auteur de ce roman historique fait naître chez Julie un sentiment d'admiration.*
 - c. *Julie a beaucoup d'admiration envers l'auteur de ce roman historique.*

La compétence paraphrastique (CP) d'un locuteur est caractérisée par Milićević (2007b : 1) comme *capacité de produire des phrases synonymes ou quasi-synonymes*. Comme souligné par Fuchs (1982 : 93), cette compétence fait partie intégrale de la compétence linguistique générale et témoigne de l'aisance langagière d'un locuteur ; « s'entraîner à paraphraser, c'est d'abord apprendre à jouer avec la diversité des formes d'expression. La capacité à paraphraser constitue l'un des tests classiques de la maîtrise de la langue » (Fuchs 1994 : 12). Selon Mel'čuk (1997 : 17), « la compétence linguistique du locuteur consiste, avant tout, en sa capacité de produire, pour un sens de départ, tous les textes qui peuvent l'exprimer (= toutes les paraphrases possibles) et de choisir le ou les textes les mieux adaptés à une situation ou à un contexte donnés ». Martinot (2010 : 63) estime que l'acquisition de la compétence paraphrastique contribue à l'apprentissage de la langue maternelle, en jouant un rôle essentiel dans le développement de la complexité de la langue d'un enfant.

Pourquoi est-il important de savoir paraphraser ? Tout d'abord, la capacité de reformuler son propos est cruciale dans la production langagière. Ainsi, dans la communication orale, un locuteur peut apporter plus de précision à son discours, le rendre plus clair, en le reformulant. Si une lexie lui échappe, il peut la remplacer par une autre, sémantiquement équivalente, sans interrompre son message. Ensuite, faire preuve d'une bonne compétence paraphrastique est important dans l'écriture académique, où souvent il faut faire une synthèse des documents en reformulant les idées des autres en ses propres mots. De plus, la CP joue un rôle important dans l'analyse d'un message. Afin de mieux comprendre l'interlocuteur, on utilise souvent la stratégie de reformulation de son énoncé. Finalement, la CP contribue à la compréhension des textes compliqués, car après avoir reformulé un texte, un locuteur parvient à mieux saisir son sens (Hsia 1991, Daunay 2004).

La CP est de première nécessité pour exercer plusieurs métiers liés à la rédaction, tels que réviseur, rédacteur ou traducteur. Par exemple, les traducteurs doivent témoigner d'une bonne compétence paraphrastique. D'une part, il leur faut garder le sens de la phrase source, « en mettant en équation des éléments, segments ou fragments du texte source » (Ballard, 2004 : 59). D'autre part, ils sont censés utiliser les moyens linguistiques appropriés de la langue d'arrivée pour exprimer d'une façon idiomatique le sens de départ, tout en appliquant les connaissances discursives qui assurent le transfert du « vouloir dire » du texte de départ.

Quelles connaissances indispensables faut-il acquérir pour produire des phrases synonymes ? Un locuteur a non seulement besoin de plusieurs types de connaissances linguistiques, mais aussi de connaissances

extralinguistiques. D'une part, il doit connaître les lexies d'une langue et les utiliser correctement dans le discours, en exploitant de nombreux liens lexicaux, tels que la synonymie, l'antonymie, la dérivation, l'intensification, etc. D'autre part, il lui est impossible de produire des paraphrases correctes sans connaissances en syntaxe et en morphologie. Il faudrait également tenir compte des connaissances discursives, qui assurent le passage de la langue au discours. Ces connaissances permettent de choisir la bonne variante paraphrastique parmi plusieurs variantes possibles dans une situation de communication particulière. Ainsi, le locuteur doit se baser sur différences dialectales, différences de registres, formules de politesse, et, entre autres, règles de cohésion textuelle. Outre les connaissances lexicales, grammaticales et discursives, la compétence paraphrastique englobe aussi la maîtrise des règles de paraphrasage qui décrivent les liens d'équivalence entre les expressions linguistiques (Mel'čuk 1992 : 9-58, Milićević 2007a : 163-331). Enfin, les connaissances extralinguistiques sont nécessaires pour établir certaines équivalences sémantiques entre la phrase de départ et la phrase d'arrivée. Ces connaissances aident à établir des équivalences non prévues par la langue.

Nous nous intéressons à la production langagière de l'apprenant avancé de français langue seconde (L2) dont les caractéristiques correspondent à celles décrites dans Bartning (1997 : 10). Un tel apprenant a étudié la L2 en tant que langue étrangère à l'école et il continue ensuite à l'étudier à l'université. Il a reçu de l'input par l'enseignement; cet apprenant guidé possède des connaissances métalinguistiques de la L2 et son but est de devenir bilingue afin d'exercer des activités professionnelles. Malgré ses compétences langagières développées, cet apprenant n'a pas encore atteint l'aisance langagière d'un locuteur natif. Il s'agit d'un locuteur indépendant, qui sait s'exprimer en L2, mais qui commet encore des erreurs dans la production orale et écrite.

L'apprenant en question possède des connaissances grammaticales assez développées (Bartning 1997, Bartning & Schlyter 2004), mais il a besoin d'améliorer ses connaissances lexicales comprenant celles des collocations (Nesselhauf 2004, 2005, Howarth 1998, Granger 1998, Lesniewska 2006, Bolly 2011). De plus, un apprenant avancé éprouve des difficultés pour reformuler un énoncé (Keck 2006, McInnis 2009, Milićević & Tsedryk 2011). Ce manque caractérise non seulement des étudiants L2 du premier cycle, mais également des étudiants du deuxième et du troisième cycles d'études (Pecorati 2003, 2008). Par ailleurs, les apprenants L2 trouvent la tâche de paraphrasage (production des paraphrases) difficile et y mettent deux fois plus de temps que les locuteurs natifs (McInnis 2009). Sun (2009) souligne que des étudiants de cycles supérieurs manquent de connaissances sur les stratégies paraphrastiques.

Notons qu'il y a un manque évident de recherches sur le processus de paraphrasage en français; nous trouvons plus de références bibliographiques à ce sujet dans la littérature sur l'anglais L2. Plusieurs de ces études (Hsja 1991, Keck 2006, Pecorati 2003, Shi 2004, McInnis 2009) ont cherché à établir un lien entre la compétence paraphrastique des apprenants L2 et le problème du plagiat dans l'écriture académique, en supposant qu'un étudiant ayant des difficultés à reformuler les idées des autres, serait enclin à recopier directement plusieurs parties d'un texte source et ainsi plagier, d'une façon non intentionnelle. Par exemple, selon Keck (2006), dans une tâche de production d'un résumé, un étudiant avancé L2 recopie presque deux fois plus de mots du texte source qu'un locuteur natif, qui paraphrase plus « en ses propres mots ».

Bien que les études en anglais L2 nous renseignent sur l'insuffisance des moyens paraphrastiques chez l'apprenant L2, elles ne reflètent qu'un seul critère dans l'évaluation de la compétence paraphrastique, à savoir, l'acceptabilité de la paraphrase du point de vue formel, et plus précisément l'analyse des paraphrases selon le nombre de mots copiés du texte source. Ces études ne décrivent pas d'une façon détaillée les moyens paraphrastiques utilisés par les participants. Or, ces informations sont particulièrement précieuses si l'on veut découvrir des lacunes de l'apprenant avancé quant au paraphrasage et par la suite proposer des moyens didactiques pour y remédier.

L'objectif de notre recherche est d'établir des différences et des similitudes entre les locuteurs natifs et les apprenants avancés de français dans la production des paraphrases.¹ Plus précisément, nous voulons répondre aux questions suivantes :

- Est-ce que le concept de paraphrase représente une difficulté particulière aux apprenants de L2 ? Autrement dit, est-ce que les apprenants de L2 et les locuteurs natifs sont significativement différents quant à la préservation du lien paraphrastique ?
- Quels types d'opérations paraphrastiques sont les plus et les moins utilisés par les deux groupes lors d'une tâche de paraphrasage ?

2 Évaluation de la compétence paraphrastique

Le cadre théorique adopté dans cette étude est la Théorie Sens-Texte (TST, Mel'čuk 1997, Kahane 2003). Cette théorie propose pour une langue donnée un modèle Sens-Texte décrivant les règles qui effectuent le passage d'un *sens* qu'un locuteur veut exprimer vers sa réalisation linguistique, *le texte*. Le modèle Sens-Texte est centré sur la production langagière et possède des outils formels de la description du paraphrasage, tels que fonctions lexicales (Žolkovkij & Mel'čuk 1966, Mel'čuk 1996, Wanner, éd. 1996) qui servent à modéliser des relations lexicales et des règles de paraphrasage. Puisque ce modèle est basé sur des concepts universellement applicables à des langues différentes, il est possible de décrire et de comparer le paraphrasage et les moyens paraphrastiques utilisés en L1 et en L2 des locuteurs.

Selon notre hypothèse, la compétence paraphrastique des locuteurs natifs est meilleure que celle des apprenants L2, en ce qui concerne l'utilisation de types variés d'opérations paraphrastiques. Nous voulons identifier les différences significatives entre les deux groupes en vue de cerner les difficultés particulières des apprenants avancés du français L2. Dans les sections qui suivent, nous décrivons la méthodologie et les résultats obtenus.

2.1 Méthodologie

2.1.1 Participants

Les participants de l'étude sont des apprenants du français L2 (n=40) de troisième et de quatrième années d'études du premier cycle des universités canadiennes. Ces apprenants ont passé le test de compétence linguistique élaboré par Tremblay (2011). Ce test à trous a été conçu afin de distinguer des niveaux de compétence linguistique en français L2. Il nous a permis d'identifier des apprenants ayant approximativement le même niveau de compétence langagière. De 65 participants, nous avons choisi 40 étudiants les plus forts (*Intermediate-High* et *Advanced* en termes de Tremblay).

Les participants L2 de l'étude ont en moyenne 22,3 ans ; il y a plus de femmes que d'hommes (32 femmes et 8 hommes). Le groupe témoin (locuteurs natifs) se compose de 40 jeunes adultes d'origine québécoise dont la moyenne d'âge est de 19,2 ans. Comme pour les apprenants de L2, il y a plus de femmes que d'hommes (26 femmes et 14 hommes) dans ce groupe.²

2.1.2 Test

Tous les participants ont écrit le test de paraphrasage dans lequel il a fallu reformuler 5 phrases proposées, en donnant trois reformulations différentes par phrase source. Presque chaque phrase à reformuler était accompagnée d'une courte phrase d'introduction expliquant le contexte et permettant de préciser son sens. Les locuteurs n'ont pas obtenu de formation explicite sur les stratégies de reformulation préalablement. Le concept de paraphrase leur a été expliqué brièvement avant le test.

En (2) sont présentées les phrases à reformuler (le contexte est précisé entre parenthèses).³ Bien que les phrases à reformuler soient assez simples structurellement et lexicalement, elles se prêtent à des reformulations variées.

- (2) a. [Aujourd'hui dans le cours de grammaire les étudiants ont travaillé sur la structure syntaxique de la phrase.] *Natalie a correctement analysé cette phrase difficile.*

- b. [Pendant le cours de français nous apprenons beaucoup d'expressions différentes.] *Marc a toujours beaucoup de questions pour le professeur de français.*
- c. *Julie admire énormément l'auteur de ce roman historique.*
- d. [Chaque samedi Paul va au théâtre pour découvrir de nouveaux acteurs.] *Paul a beaucoup d'enthousiasme pour le théâtre.*
- e. [L'étudiant ne savait pas quels cours il devait suivre à l'automne prochain.] *Le professeur a donné un bon conseil à l'étudiant.*

2.1.3 Codage et analyse des données

Les données obtenues ont été transcrites et analysées avec le logiciel *Statistical Package for the Social Sciences* (SPSS). L'analyse s'est faite sur deux axes : (i) préservation du lien paraphrastique et (ii) opérations paraphrastiques effectuées. Selon la préservation du lien paraphrastique, les phrases obtenues ont été divisées en paraphrases (P), paraphrases défectueuses ([?]P) et non-paraphrases (nonP).

Par exemple, les phrases en (3) sont considérées comme des paraphrases puisqu'elles ont un sens équivalent.

- (3) a. *Le professeur a donné un bon conseil à l'étudiant.*
- b. *Le professeur a bien conseillé l'étudiant.*

La phrase (4b) est considérée comme paraphrase défectueuse, car elle garde le sens de la phrase de départ (4a), mais contient une erreur lexicale : dans (4b) le locuteur a utilisé une substitution synonymique incorrecte en remplaçant *l'auteur de ce roman* par [?]*l'écrivain de ce livre*.

- (4) a. *Julie admire énormément l'auteur de ce roman historique.*
- b. *Julie admire [?]l'écrivain de ce livre à contenu historique.*

Les phrases qui n'ont pas conservé le sens de la phrase de départ ont été classées comme non-paraphrases. Par exemple, la phrase (5b) n'est pas sémantiquement équivalente à la phrase (5a). Nous ne les considérons pas comme des paraphrases.

- (5) a. *Le professeur a donné un bon conseil à son étudiant.*
- b. *L'étudiant a tenu compte de l'avis de son professeur.*

Les paraphrases classées comme P et [?]P ont été ensuite analysées selon les types d'opérations paraphrastiques effectuées. Cette analyse s'est basée sur la typologie standard de paraphrases de la Théorie Sens-Texte (Milićević 2007a). Les opérations paraphrastiques sont décrites au moyen de règles de paraphrasage qui établissent l'équivalence entre la phrase de départ et la phrase d'arrivée. Par exemple, les règles de paraphrasage lexico-syntaxiques reposent sur de nombreuses relations lexicales qui existent entre les unités lexicales et sont fréquemment appliquées par les locuteurs. En particulier, une des règles lexico-syntaxiques utilisée en français de façon récurrente est la suivante : un verbe biactanciel peut être remplacé par une construction constituée d'une nominalisation de ce verbe et du verbe support correspondant (ex. *décider* ~ *prendre une décision*, *accuser* ~ *porter des accusations*, *respecter* ~ *éprouver du respect*). À titre d'illustration, en (6), considérons quelques relations lexicales de la lexie ADMIRATION ('admiration de l'individu X devant Y pour Z').

- (6) Terme générique pour L : *sentiment d'admiration*
Nom pour X : *admirateur de*, *fan fan*
Nom pour Z : *source [d'admiration]*, *objet de l'admiration*
Adjectif pour Y : *admirable*
Y est l'objet de l'admiration de X : *faire l'admiration de N_X*, *être l'objet de l'admiration*
X éprouve de l'admiration devant Y (= verbe) : *X admire Y*, *X adore Y*, *X apprécie beaucoup Y*

En mettant en valeur ces relations lexicales et en appliquant des règles lexico-syntaxiques de paraphrasage, il est possible de produire des paraphrases en (7) pour la phrase de départ *Julie admire énormément l'auteur de ce roman historique*. Nous n'entrons pas dans les détails de la description formelle des règles de paraphrasage adoptée dans la Théorie Sens-Texte ici ; il est possible de trouver plus d'information à ce sujet dans Milićević (2007a).

- (7) *Julie éprouve un sentiment d'admiration pour l'auteur de ce roman historique.*
Julie est une grande admiratrice de l'auteur de ce roman historique.
L'auteur de ce roman historique est une grande source d'admiration pour Julie.
Pour Julie, l'auteur de ce roman historique est tout simplement admirable.
L'auteur de ce roman historique fait l'admiration de Julie.
Julie apprécie beaucoup l'auteur de ce roman historique.
Julie adore l'auteur de ce roman historique.

Pour faciliter l'analyse des données, nous avons adopté un système de points attribuables à chaque opération paraphrastique produite. Les opérations se divisent en quatre types majeurs : les opérations linguistiques lexico-syntaxiques (LS), syntaxiques (Synt) et sémantiques (Sém), ainsi que les opérations extralinguistiques (Ext). Nous illustrons le codage des données dans ce qui suit.

Les opérations lexico-syntaxiques incluent différents types de substitutions lexicales, accompagnées quelquefois de restructurations syntaxiques. Les substitutions peuvent être simples (quand une lexie est remplacée par une autre lexie, *des informations ~ des renseignements*) et complexes (par exemple, quand on remplace une lexie par une construction avec un verbe support, *informer ~ transmettre l'information*).

- (8) a. *Julie admire l'auteur de ce roman historique. ~ Julie adore l'auteur de ce roman historique.*
b. *Le professeur a donné un bon conseil à l'étudiant. ~ L'enseignant a offert un judicieux conseil à son disciple.*

Par exemple, pour la substitution simple *admirer ~ adorer* nous attribuons un point LS à la paraphrase en (8a). En (8b), le locuteur a utilisé quatre substitutions synonymiques : *professeur ~ enseignant, étudiant ~ disciple, bon [conseil] ~ judicieux [conseil]* et *donner [un conseil] ~ offrir [un conseil]*. Nous attribuons 4 points LS à cette paraphrase.

Les opérations sémantiques constituent des modifications du sens de la phrase de départ ou des remplacements d'une lexie par sa définition lexicographique. Par exemple, l'opération sémantique d'inférence est illustrée en (9a). Le lien d'inférence est 'X donne un bon conseil à Y' ~ 'X aide Y par un conseil'.

- (9) a. *Le professeur a donné un bon conseil à l'étudiant. ~*
Le professeur a aidé l'étudiant par ses conseils.
b. *l'auteur [de ce roman] ~ la personne qui a écrit [ce roman]*

Dans (9b) il s'agit de la substitution du signifié d'une lexie L par sa décomposition sémantique. Ainsi, 'X est l'auteur de Y' a été remplacé par 'X est la personne qui a écrit Y'. Appliquée à la phrase *Julie admire l'auteur de ce roman*, cette opération produit une paraphrase exactement équivalente *Julie admire la personne qui a écrit ce roman*. Nous attribuons un point Sém pour chacune des substitutions illustrées en (9).

Les opérations syntaxiques sont celles qui s'appliquent à la structure syntaxique de la phrase. Par exemple, les opérations syntaxiques de passivisation et de relativisation appliquées à la phrase (2a) sont respectivement illustrées en (10a) et (10b). Nous attribuons un point Synt à chaque paraphrase.

- (10) a. *Natalie a correctement analysé cette phrase difficile. ~ Cette phrase difficile a été bien analysée par Natalie.*
b. *Natalie a correctement analysé cette phrase difficile. ~ La phrase que Natalie a bien analysée était difficile.*

Les paraphrases extralinguistiques sont très approximatives et sont produites quand un locuteur utilise des connaissances sur la situation de communication, des connaissances générales sur le monde et ses capacités logiques. Par exemple, en (11a) il s'agit des connaissances encyclopédiques (pour remplacer 'français' par 'langue de Molière' il faut savoir qui était Molière). En (11b) un locuteur a effectué une opération extralinguistique ayant ajouté de l'information 'sur des expressions françaises' retrouvée dans la phrase de contexte.⁴ Nous attribuons 1 point Ext pour chacune des opérations en (11).

- (11) a. [professeur] *de français* ~ [professeur] *de la langue de Molière*
b. *questions au professeur de français* ~ *questions sur des expressions françaises*

Souvent, pour produire une paraphrase, un locuteur effectue plusieurs opérations paraphrastiques en même temps.

- (12) *Marc a toujours beaucoup de questions pour le professeur de français.* ~
Le professeur de français répond aux nombreuses questions que Marc lui pose.

Ainsi, la paraphrase en (12) est le résultat de différents types d'opérations paraphrastiques. La substitution *avoir des questions* ~ *répondre aux questions* est une inférence sémantique : celui qui a des questions s'attend habituellement à recevoir une réponse, quoique ce ne soit pas toujours le cas. La phrase contient aussi deux substitutions lexicales *avoir* [des questions] ~ *poser* [des questions] et *beaucoup* [de questions] ~ *nombreuses* [questions]. Ces substitutions sont accompagnées de deux opérations syntaxiques indépendantes ; ce sont la relativisation et la pronominalisation du complément d'objet indirect : *des questions de Marc* [au professeur] ~ *des questions que Marc lui* [= au professeur] *pose*. Nous attribuons donc 1 point Sém, 2 points LS et 2 points Synt à cette phrase.

2.2 Résultats

En ce qui concerne le test de paraphrasage, il s'agit de présenter les résultats du test statistique ANOVA (*l'analyse statistique de variance à un facteur*). Nous utilisons ce test afin de comparer les moyennes pour les groupes étudiés. Notre objectif est de vérifier les critères selon lesquels la production des paraphrases pour les deux groupes serait significativement différente. Les résultats des tests nous indiquent dans quelle mesure la performance des locuteurs natifs dans la production des paraphrases est meilleure de celle des apprenants L2. Les variables dépendantes sont (i) le nombre de paraphrases produites et (ii) les types des opérations paraphrastiques utilisées. Notre objectif est de vérifier si ces variables dépendantes sont en relation avec la variable indépendante, le groupe langagier (L1 vs L2). Nous choisissons le seuil de signification (risque d'erreur) de 5 %. La valeur de F critique dans une table de Snedecor avec le seuil de probabilité de 0,05 pour le degré de liberté F (1, 78) est égale à 4,00.

2.2.1 Le concept de paraphrase

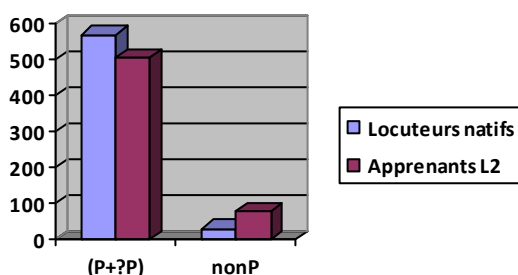
Notre première question de recherche était « Est-ce que les apprenants de L2 et les locuteurs natifs sont significativement différents quant à la préservation du lien paraphrastique ? »

Nous avons calculé le nombre de paraphrases en additionnant les paraphrases et les paraphrases défectueuses (P +²P) produites par participant afin de déterminer si le concept de paraphrases est clair pour les participants des deux groupes. Ces données sont présentées dans le Tableau 1.

	Nombre de (P+ ² P)	Nombre de NonP	Total
Locuteurs natifs	570	30	600
Apprenants L2	507	76	583

Tableau 1 Paraphrases et de non-paraphrases produites par les participants

Rappelons que nous avons demandé aux participants de proposer trois reformulations par phrase source, ce que les locuteurs natifs ont pu faire. Le corpus L1 comprend 600 phrases. En revanche, certains apprenants L2 n'ont pas été en mesure de produire 3 reformulations différentes; ainsi, le corpus L2 contient en tout 583 phrases. Nous pouvons observer que la plupart des phrases proposées par les L1 et par les L2 sont des phrases sémantiquement équivalentes. Pour l'analyse statistique, la variable dépendante est le nombre de paraphrases produites (P+ ?P). Le test ANOVA montre qu'il y a une différence significative entre les groupes ($F(1, 78) = 13,239, p < ,001$) en ce qui concerne la production des paraphrases vs non-paraphrases. Le Graphique 1 représente les pourcentages des paraphrases et des non-paraphrases dans les deux corpus.



Graphique 1 Taux de paraphrases et de non-paraphrases dans les corpus L1 et L2

Afin de vérifier si la différence entre les groupes est grande, nous avons calculé la taille de l'effet (Cohen 1988) η^2 pour ANOVA. La statistique η^2 (η^2) est calculée à partir des éléments déjà connus : il faut diviser la somme des carrés intergroupe par la somme totale des carrés. On obtient la valeur de η^2 , qui s'interprète de la manière indiquée dans le Tableau 2 (Kinnear & Gray 2008 : 259).

Taille de l'effet (d)	η^2	Taille de l'effet
$0,2 \leq d < 0,5$	$0,1 \leq \eta^2 < ,06$	Effet de petite taille
$0,5 \leq d < 0,8$	$,06 \leq \eta^2 < ,14$	Effet de taille moyenne
$d \geq 0,8$	$\eta^2 \geq ,14$	Effet de grande taille

Tableau 2 Taille de l'effet de Cohen (1988)

Pour la première mesure, soit la production des paraphrases vs non-paraphrases, nous observons la différence significative entre les groupes avec l'effet de taille moyenne ($\eta^2=0,13$).

2.2.2 L'utilisation d'opérations paraphrastiques différentes

Notre deuxième question de recherche était «Quels types d'opérations paraphrastiques sont les plus et les moins utilisés par les deux groupes lors d'une tâche de paraphrasage ?» Cette question a quatre sous-questions. Nous avons étudié l'utilisation de 4 types majeurs d'opérations paraphrastiques : opérations utilisées pour produire des paraphrases extralinguistiques, des paraphrases sémantiques, des paraphrases lexico-syntaxiques et des paraphrases syntaxiques.

Opérations utilisées pour produire des paraphrases extralinguistiques

La variable dépendante est le nombre d'opérations paraphrastiques utilisées pour produire des paraphrases extralinguistiques par participant. Les résultats du test ANOVA montrent qu'il n'y a pas de différence significative entre les groupes ($F(1, 78) = 3,113, p = ,078$). La valeur de p étant largement supérieure au seuil de signification 0,05 et la valeur de F étant moindre que la valeur critique, nous

concluons que les deux groupes ne diffèrent pas significativement en ce qui concerne l'utilisation des opérations extralinguistiques.

Opérations utilisées pour produire des paraphrases sémantiques

La variable dépendante est le nombre d'opérations paraphrastiques utilisées pour produire des paraphrases sémantiques. Les deux groupes diffèrent significativement ($F(1, 78) = 12,098, p < ,005, \eta^2 = 0,13$), avec l'effet de taille moyenne.

Opérations utilisées pour produire des paraphrases lexico-syntaxiques

La variable dépendante est le nombre d'opérations lexico-syntaxiques effectuées par participant. Les résultats du test montrent une différence très significative entre les deux groupes ($F(1, 78) = 38,149, p < ,001, \eta^2 = 0,33$). Les deux groupes diffèrent significativement avec l'effet de grande taille en ce qui concerne l'utilisation des opérations lexico-syntaxiques.

Opérations utilisées pour produire des paraphrases syntaxiques

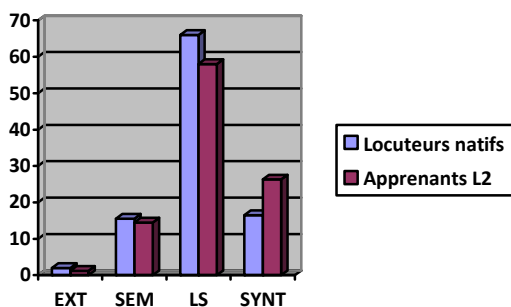
La variable dépendante est le nombre d'opérations syntaxiques effectuées par participant. Le test a montré une différence non significative entre les moyennes des groupes L1 et L2 ($F(1, 78) = ,017, p = ,896$).

Nous avons relevé 1510 occurrences des opérations paraphrastiques. Les francophones en ont effectué beaucoup plus (n=925) que les apprenants du français L2 (n=585). Cela concerne surtout les opérations lexico-syntaxiques. Ces données sont présentées dans le Tableau 3.

Type d'opération paraphrastique	Opérations extralinguistiques	Opérations linguistiques			Total
		Sém	Lex-Synt	Synt	
Locuteurs natifs	19	144	607	155	925
Apprenants L2	7	85	338	155	585

Tableau 3 Opérations paraphrastiques effectuées par les locuteurs natifs et les apprenants L2

Les pourcentages de différents types d'opérations effectuées par les participants de deux groupes sont illustrés dans le Graphique 2. En examinant les taux de chaque type d'opérations paraphrastiques, nous remarquons que les deux groupes ont utilisé tous les types d'opérations paraphrastiques, en donnant une nette préférence aux opérations lexico-syntaxiques. Les locuteurs natifs ont légèrement devancé les apprenants L2 dans l'utilisation de tous les types d'opérations paraphrastiques à l'exception du type syntaxique. Les apprenants du français L2 ont utilisé plus souvent les opérations syntaxiques (26,3 % vs 16,5 %) dans le paraphrasage.



Graphique 2 Les taux d'utilisation de différents types d'opérations paraphrastiques : L1 vs L2

3 Discussion

Les résultats de notre étude confirment notre hypothèse de départ, d'après laquelle la compétence paraphrastique des locuteurs natifs est meilleure que celle des apprenants du français L2. Nous avons appris plusieurs faits sur les particularités de la CP chez les locuteurs natifs et les apprenants du français L2. Notre première question de recherche concernait la préservation du lien paraphrastique. Les deux groupes testés semblent ne pas avoir de difficultés à produire des paraphrases. Dans ce test, il s'agissait des phrases structurellement et lexicalement assez simples, pour lesquelles il y avait beaucoup de reformulations potentielles. Le nombre de phrases sémantiquement équivalentes a été élevé dans les deux corpus (95 % dans le corpus L1 et 87 % dans le corpus L2). Nous pouvons conclure que les apprenants du français L2 avancés, en général, savent produire des paraphrases. Néanmoins, le test statistique ANOVA a montré que, pour le critère de production des paraphrases par rapport aux non-paraphrases, les groupes L1 et L2 se distinguent significativement avec l'effet de taille moyenne. La production des locuteurs non natifs contient deux fois plus de non-paraphrases que celle des locuteurs natifs (13 % vs 5 %). Par conséquent, il serait bénéfique pour les apprenants du français L2 avancés de recevoir une formation explicite sur le concept de paraphrase.

Notre deuxième question de recherche était liée aux types d'opérations paraphrastiques effectuées par les participants de deux groupes. Nous voulions savoir si les participants connaissaient les différents types de paraphrases et si un des groupes favoriserait le recours à des opérations paraphrastiques d'un type particulier. L'analyse des résultats montre que les participants des deux groupes utilisent tous les types d'opérations paraphrastiques. Les paraphrases lexico-syntaxiques sont les plus fréquentes dans les deux corpus. Le nombre d'opérations lexico-syntaxiques appliquées est plus grand dans le corpus des locuteurs natifs que dans celui des apprenants L2. Le taux des opérations impliquées dans la production des paraphrases extralinguistiques est très faible dans les deux groupes.

Les tests statistiques ANOVA ont montré une différence non significative pour le type de paraphrases extralinguistiques et syntaxiques, une différence significative (effet de taille moyenne) pour le type sémantique et une différence très significative (effet de grande taille) pour le type lexico-syntaxique. Ainsi, la difficulté de reformuler le discours des apprenants du français L2 s'explique par l'insuffisance des connaissances pour produire des paraphrases sémantiques et surtout des paraphrases lexico-syntaxiques.

Nous avons constaté que les locuteurs natifs possèdent un vocabulaire plus riche et utilisent souvent des moyens lexico-syntaxiques de paraphrasage complexes. Par exemple, pour paraphraser *X a beaucoup de questions pour Y* ils utilisent d'abord la substitution lexicale *X bombarde Y de questions* et ensuite la passivisation : *Y est bombardé de questions de X*. Notons que les paraphrases de ce type sont rares dans le corpus des apprenants du français L2, ce qui n'est pas inattendu. Pour effectuer des opérations lexico-syntaxiques complexes, il faut non seulement posséder un vocabulaire riche, mais aussi savoir bien manipuler les éléments de la phrase du point de vue syntaxique. Les locuteurs natifs restructurent les phrases en recourant à des moyens syntaxiques plus variés que les apprenants L2.

La performance des locuteurs natifs est nettement supérieure dans le cas des paraphrases lexico-syntaxiques. En moyenne, les francophones ont proposé à peu près deux fois plus de variantes différentes de substitutions lexicales par participant que les apprenants du français L2. Les locuteurs natifs se sont servis plus de mots rares que les locuteurs non natifs; la variabilité de moyens lexicaux qu'ils utilisent dans le paraphrasage est plus élevée.

Didactique de la paraphrase

Afin d'améliorer la capacité de paraphraser des apprenants avancés, il faut faire appel à l'enseignement explicite des notions reliées au paraphrasage, telles que paraphrase, opération paraphrastique, règle de paraphrasage, types de paraphrases et relations lexicales. Voici quelques principes qui nous semblent essentiels dans l'enseignement des notions reliées à la paraphrase:

- Étant donné que les apprenants L2 utilisent rarement des opérations de paraphrasage sémantiques et lexico-syntaxiques, il faut les leur enseigner explicitement.

- Le transfert négatif de la langue maternelle qui se manifeste encore chez certains apprenants doit être pris en considération dans la didactique de la paraphrase ; les exercices contrastifs en L1 et L2 peuvent aider à éviter les erreurs dues au transfert.
- Puisque les apprenants font souvent des erreurs de collocations, surtout avec des verbes supports, il est nécessaire d'enseigner explicitement des collocations pour faciliter la production de paraphrases correctes. Il faut également mettre en évidence les erreurs commises dans les décompositions sémantiques.
- Les exercices sur la paraphrase doivent non seulement inclure des lexies de haute fréquence, mais aussi des lexies de moindre fréquence.

Les formalismes de la Théorie Sens-Texte pourraient être utilisés dans l'enseignement du français L2 aux niveaux avancés. Cependant, il est important de rendre ces formalismes accessibles à l'apprenant en utilisant le langage naturel et le métalangage des grammaires traditionnelles (ex. *sujet, complément d'objet direct, nom, verbe, adverbe* etc.). Le Tableau 4 illustre quelques correspondances entre les formalismes de la TST et ceux que nous proposons dans l'explication des règles de paraphrasage.

Règle « classique » de la TST	Règle conviviale	Exemple
$L(V) \approx \text{Oper}_1(S_{0(L)}) - II \rightarrow S_0(L_{(V)})$	$x V_Y \approx V_{\text{sup}} - \text{objet direct} \rightarrow N_{(\text{Nominalisation du V})}$	<i>Il a crié. ≈ Il a poussé un cri.</i>
$L(V) \approx \text{Oper}_1(S_{1(L)}) - II \rightarrow S_1(L_{(V)})$	$x V_Y \approx \text{être} + \text{Nom pour X}$ [Pour vous aider : Comment appelle-t-on X ? X est _____ (Nom).]	<i>Charles craint tout ! ≈ fam Charles est un froussard !</i>
$L(V) \approx \text{Oper}_1(S_{2(L)}) - II \rightarrow S_2(L_{(V)})$	$x V_Y \approx \text{être} + \text{Nom pour Y}$ [Pour vous aider : Comment appelle-t-on Y ? Y est _____ (Nom).]	<i>Cette femme se passionne pour le théâtre. ≈ Le théâtre est la passion de cette femme.</i>

Tableau 4 Exemple de présentation de règles de paraphrasage lexico-syntaxiques

Puisque la maîtrise de la paraphrase contribue à la maîtrise de la langue, l'enseignement des stratégies de paraphrasage ne peut que bénéficier aux apprenants. Or, pour pouvoir enseigner ces stratégies aux élèves, l'enseignant lui-même doit posséder les connaissances nécessaires, y compris celles des règles de paraphrasage. Cependant, comme l'observe Keck (2006 : 262) à l'égard de l'anglais L2, il y a un manque de ces connaissances chez les enseignants : « English and ESL instructors need to learn methods which help students to move beyond a reliance on copying from source texts ». Nous constatons aussi ce manque dans l'enseignement du français L2.

Dans cet article, nous avons décrit des différences et des similitudes entre les locuteurs natifs et les apprenants du français L2 quant à la compétence paraphrastique. Dans la tâche de reformulation, les deux groupes ont produit un nombre élevé de paraphrases. Cependant, les apprenants L2 ont proposé plus de phrases sémantiquement non équivalentes que les locuteurs natifs. Ceci suggère que le concept de paraphrase doit être introduit explicitement aux apprenants L2. Nous avons également étudié les stratégies de reformulation utilisées par les participants de notre étude. Les deux groupes ont utilisé tous les types de paraphrases, mais dans une mesure différente. Les apprenants L2 ont utilisé moins d'opérations sémantiques et lexico-syntaxiques que les locuteurs natifs. La connaissance des relations lexicales entre les lexies d'une langue, un élément crucial pour la production des paraphrases, semble être une lacune importante chez l'apprenant avancé. Afin de développer sa compétence paraphrastique, il convient d'enseigner d'une façon explicite les notions liées au paraphrasage au niveau des études universitaires.

Références bibliographiques

- Ballard, M. (2004). La théorisation comme structuration de l'action du traducteur. *La linguistique*, 1(40), 51-66.
- Bartning, I. (1997). L'apprenant dit avancé et son acquisition d'une langue étrangère – Tour d'horizon et esquisse d'une caractérisation de la variété avancé. *AILE*, 9, 9-50.
- Bartning, I. & Schlyter, S. (2004). Itinéraires acquisitionnels et stades de développement en français L2. *French Language Studies*, 14, 281-299.
- Bolly, C. (2011). *Phraséologie et collocations. Approche sur corpus en français L1 et L2*. Bern: Peter Lang.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Lawrence Erlbaum Associates.
- Daunay, B. (2004). Réécriture et paraphrase. Contribution à une histoire des pratiques d'écriture scolaire. *Le français aujourd'hui*, 1(144), 25-32.
- Fuchs C. (1982). *La Paraphrase*. Paris : PUF.
- Fuchs, C. (1994). *Paraphrase et énonciation*. Paris : Éditions Ophrys.
- Granger, S. (1998). Prefabricated patterns in advanced EFL writing; Collocations and formulae. In *Phraseology: Theory, Analysis, and Applications*, sous la dir. de A.P. Cowie, 145-160. Oxford: Clarendon Press.
- Hsia, S. (1991). Grammaticality judgments, paraphrase and reading comprehension: evidence from European, Latin American, Japanese and Korean ESL learners. *Perspectives*, 3, 81-96.
- Howarth, P. (1998). The phraseology of learner's academic writing. In *Phraseology. Theory, Analysis and Applications*, sous la dir. de A.P. Cowie, 161-186. Oxford: Oxford University Press.
- Kahane, S. (2003). The Meaning-Text Theory. In *Dependency and Valency. An International Handbook of Contemporary Research*, vol. 1, sous la dir. de V. Agel *et al.*, 546-570. Berlin/New York: De Gruyter.
- Keck, C. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing*, 15, 261-278.
- Kinney, P. & Gray, C. D. (2008). *SPSS 15 made simple*. New York: Psychology Press.
- Leśniewska, J. (2006). Is Cross-Linguistic Influence a Factor in Advanced EFL Learners' Use of Collocations? In *Cross-linguistic Influences in the Second Language Lexicon*, sous la dir. de J. Arabski, 144-199. Clevedon, England: Cromwell Press Ltd.
- Martinot, C. (2010). Reformulation et acquisition de la complexité linguistique. *Travaux de linguistique*, 61, 63-96.
- McInnis, L. (2009). Analyzing English L1 and L2 paraphrasing strategies through concurrent verbal report and stimulated recall protocols. Thèse de maîtrise, Université de Toronto.
- Mel'čuk, I. (1992). Paraphrase et lexique : la théorie Sens-Texte et le Dictionnaire explicatif et combinatoire. In *DEC III*, sous la dir. de Mel'čuk *et al.* 9-58. Montréal : Les Presses de l'Université de Montréal.
- Mel'čuk, I. (1996). Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon. In *Lexical Functions in Lexicography and Natural Language Processing*, sous la dir. de L. Wanner, 37-102. Amsterdam/Philadelphia: John Benjamins.
- Mel'čuk, I. (1997). *Vers une linguistique Sens-Texte : leçon inaugurale faite le vendredi 10 janvier 1997*. Paris : Collège de France.
- Milićević, J. (2007a). *La paraphrase : modélisation de la paraphrase langagière*. Bern: Peter Lang.
- Milićević, J. (2007b). Paraphrase as a Tool for Achieving Lexical Competence in L2. In *Complexity, Accuracy and Fluency in Second Language Use, Learning and Teaching*, 124-134. Actes du colloque. Brussels: Royal Belgium Academy of Sciences.
- Milićević, J. & Tsedryk, A. (2011). Assessing and Improving Paraphrasing Competence in FSL. In *Proceedings of the 5th International Conference on Meaning-Text Theory*, sous la dir. de I. Boguslavsky & L. Wanner, 175-185. Barcelone, Université Pompeu Fabra.

- Nesselhauf, N. (2004). How learner corpus analysis can contribute to language teaching: A study of support verb constructions. In *Corpora and Language Learners*, sous la dir. de G. Aston, S. Bernardini & D. Stewart, 109-124. Amsterdam, Philadelphia: John Benjamins.
- Nesselhauf, N. (2005). *Collocations in a Learner Corpus*. Amsterdam/ Philadelphia: John Benjamins.
- Pecorati, D. (2003). Good and original: Plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing*, 12, 317-345.
- Pecorati, D. (2008). *Academic Writing and Plagiarism. A linguistic analysis*. London/New York: Continuum.
- Shi, L. (2004). Textual borrowing in second-language writing. *Written Communication*, 21, 171-200.
- Sun, Y.-C. (2009). Using a two-tier test in examining Taiwan graduate students' perspectives on paraphrasing strategies. *Asia Pacific Education Review*, 10(3), 399-408.
- Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research: "Clozing" the gap. *Studies in Second Language Acquisition*, 33(3), 339-372.
- Tsedryk, A. (2013). Didactique de la paraphrase : évaluation et développement de la compétence paraphrastique chez l'apprenant de français langue seconde. Thèse de doctorat non publiée. Dalhousie University, Halifax.
- Wanner, L. (ed.) (1996). *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam/Philadelphia: Benjamins.
- Žolkovskij, A. & Mel'čuk, I. (1966). O sisteme semantičeskogo sinteza. I. Stroenie slovarja. [Sur le système de la synthèse sémantique (des textes). I: Construire un dictionnaire]. *Naučno-texničeskaja informacija*, 11, 48-55.

¹ Dans cette étude, nous avons décrit les procédés associés à la production des paraphrases. Une étude concernant la saisie des énoncés paraphrastiques par les apprenants d'une L2 et par les locuteurs natifs pourrait être intéressante du point de vue de la compréhension du discours et des mécanismes qui y sont impliqués.

² Pour une description détaillée de cette étude, consulter Tsedryk (2013). Disponible au <http://dalspace.library.dal.ca/bitstream/handle/10222/16013/Tsedryk-Alexandra-PhD-FRENCH-January-2013.pdf?sequence=1>

³ Lorsque nous avons demandé à des locuteurs natifs adultes de reformuler les cinq phrases du test, certains ont trouvé la phrase (2d) naturelle, alors que d'autres l'ont trouvée douteuse [phrase douteuse : je ne suis pas sûr que je la dirais, mais elle ne m'étonnerait guère si je l'entendais]. Après avoir analysé les paraphrases de nos participants, proposées pour cette phrase, nous avons constaté que cela n'a pas empêché les participants de comprendre le sens de la phrase de départ et d'en produire des paraphrases.

⁴ Nous avons explicitement demandé aux participants de ne pas utiliser l'information contextuelle dans les paraphrases, en les prévenant que la phrase de contexte ne sert qu'à clarifier le sens de la phrase de départ. Néanmoins, certaines personnes ont exprimé les détails contextuels dans leurs productions.